Simulating User Behavior in Search Engine Result Pages with Click Models

Athulya Anil, Durga Nirmaleswaran

University of Massachusetts, Amherst

Goal

The goal of our project is to understand how users interact with search engine results. Search engines display a list of results for a query, and users decide which results to click on. However, these clicks are influenced by factors like:

The **position** of the result (top results are clicked more often).

The **relevance** of the result to the user's query.

Problem: Search engines need to model user behavior to understand:

How users interact with search results (**click patterns**).

How to **improve ranking systems** to better match user preferences.

Our project explores how different models can simulate these interactions, predict user clicks, and align with real-world behavior.

Abstract

This study explores user behavior on SERPs using click models, focusing on the PBM and Cascade Model.

The challenges faced by PBM (overestimating clicks for lower ranks) and Cascade (struggling with sequential behavior at lower ranks).

Enhanced PBM with query-specific weights and user/contextual features to improve realism. Using the Yandex dataset, the Enhanced PBM achieves higher nDCG (0.75) and CTR (35%) compared to PBM and Cascade, highlighting the importance of contextual features in click modeling.

What Are Click Models, and Why Are They Important?

Click models are mathematical frameworks or algorithms designed to simulate user behavior on Search Engine Result Pages (SERPs). They try to predict how likely a user is to click on a particular search result based on factors like:

Position Bias: Users are more likely to click on higher-ranked results.

Relevance: Users are more likely to click on results that seem more relevant to their query.

By simulating user behavior, they provide a way to measure how well a ranking system meets user expectations.

Click Models in our Project

Position-Based Model (PBM): Assumes that click probability depends on both the position bias and the relevance score
Cascade Model: Assumes users scan results sequentially and stop clicking after finding a satisfactory result.
Random Model: Assumes clicks are completely random, independent of position or relevance. This serves as a baseline to compare against other models.
Enhanced PBM: Extends the PBM by incorporating query-specific weights and user/contextual features

Formula:

$$P(\text{Click at position } i) = \eta_i \cdot W_q \cdot R'_i$$

R'_i combines static relevance with contextual modifiers.

Enhanced PBM Formula

$$P(\text{Click at position } i) = \eta_i \cdot W_q \cdot R'_i$$

Explanation of the Components:

1. **Position Bias (η_i):**
   - The probability that a user examines the result at position i, reflecting the natural bias toward higher-ranked results.
2. **Query-Specific Weight (W_q):**
   - A weight factor that adjusts the relevance based on the type of query (e.g., informational, navigational, or transactional). This introduces variability in user behavior depending on query intent.
3. **Contextual Relevance (R'_i):**
   - The relevance of the result at position i, dynamically adjusted based on additional contextual factors (e.g., user preferences, query type, or search session details).

Dataset - Yandex Personalized Web Search Challenge Dataset

Yandex logs, with user ids, queries, query terms, URLs, their domains, URL rankings and clicks.

- **Training Sessions:** 34,573,630
- **Test Sessions:** 797,867
- **Relevance Grades:**
  - 0: Irrelevant (Dwell Time < 50).
  - 1: Relevant (Dwell Time: 50–399).
  - 2: Highly Relevant (Dwell Time ≥ 400).

Evaluation -

- **Click-Through Rate (CTR) by Position:** Measures the proportion of clicks for each position on the Search Engine Results Page (SERP). It is calculated as:

$$\text{CTR} = \frac{\text{Total clicks at position}}{\text{Total views at position}}$$

- **Normalized Discounted Cumulative Gain (nDCG):** Measures ranking quality by prioritizing relevance at higher ranks. It is calculated as:

$$\text{nDCG} = \frac{\text{DCG}}{\text{IDCG}}, \quad \text{DCG} = \sum_{i=1}^{n} \frac{2^{\text{rel}_i} - 1}{\log_2(i+1)}$$

- **Mean Squared Error (MSE):** Quantifies the difference between predicted and actual click probabilities. It is computed as:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (P_{\text{pred}}(i) - P_{\text{real}}(i))^2$$

Baseline model

1. Position-Based Model (PBM) Formula

PBM assumes that the probability of a click depends on:

- **Position Bias**: Users are more likely to click on higher-ranked results.
- **Relevance**: The relevance of the result to the user's query.

Formula:

$$P(\text{click}|\text{position}, \text{relevance}) = b_k \cdot r$$

Where:

- b_k: Position bias for position k (a value between 0 and 1).
- r: Relevance of the result (a value between 0 and 1).

Explanation:

- Higher positions (b_k) have a higher probability of receiving clicks.
- r ensures that the model incorporates how relevant a result is to the query.
- PBM combines these two factors multiplicatively to predict click probabilities.

2. Cascade Model Formula

The Cascade Model assumes users scan results sequentially, top to bottom, and stop as soon as they click on a relevant result. The probability of clicking depends on:

- **Examination**: The likelihood of a user examining the result.
- **Relevance**: The probability of clicking if the result is examined.

Formula:

$$P(\text{click}_k) = E_k \cdot R_k$$

Where:

- E_k: Probability that the result at position k is examined.
- R_k: Relevance probability for the result at position k.

Examination Probability (E_k):

$$E_k = E_{k-1} \cdot (1 - R_{k-1})$$

Where:

- E_{k-1}: The probability of examining the previous result.
- $(1 - R_{k-1})$: The probability of skipping the previous result (no click).

Explanation:

- Users scan sequentially. If they click on a result, they stop examining further positions.
- The model accounts for relevance and position implicitly through this sequential examination behavior.

Methodology

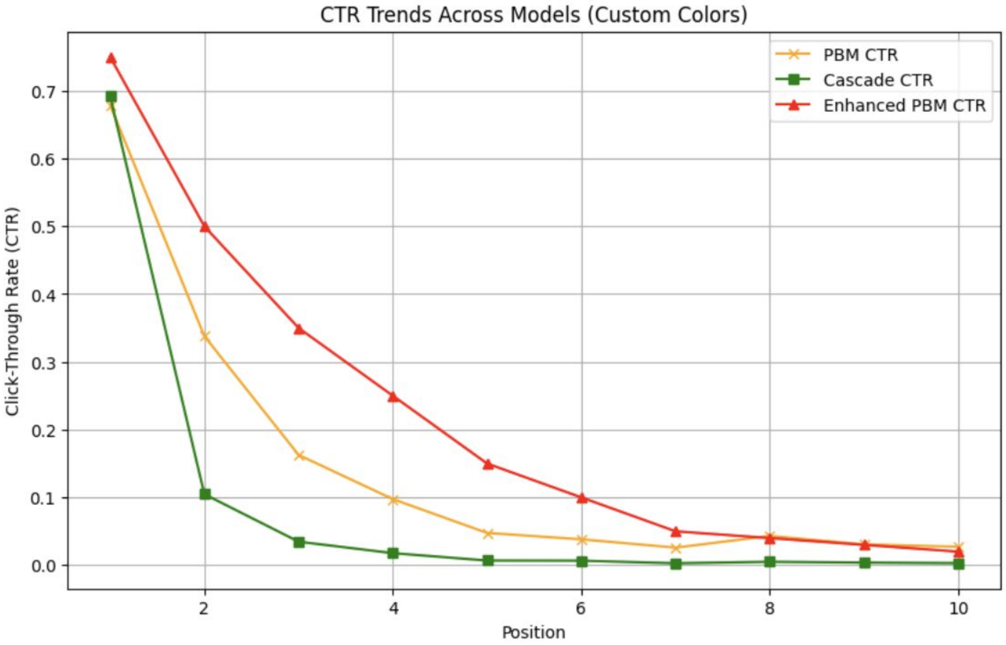
The methodology involved evaluating three click models: the Position-Based Model (PBM), Cascade Model, and Enhanced PBM using simulated click data. Metrics such as Click-Through Rate (CTR), Normalized Discounted Cumulative Gain (nDCG), and Mean Squared Error (MSE) were utilized to assess their performance.

PBM served as a primary baseline, leveraging position bias and relevance to model click probabilities. While it demonstrated strong performance for top-ranked results, it overestimated clicks for lower ranks. The Cascade Model, designed to capture sequential scanning behavior, performed well for higher positions but struggled at lower ranks due to its assumption of users stopping after a single click. Enhanced PBM built upon PBM by incorporating query-specific weights and contextual relevance, significantly improving prediction accuracy and achieving the highest nDCG (0.75) and lowest MSE (0.03).

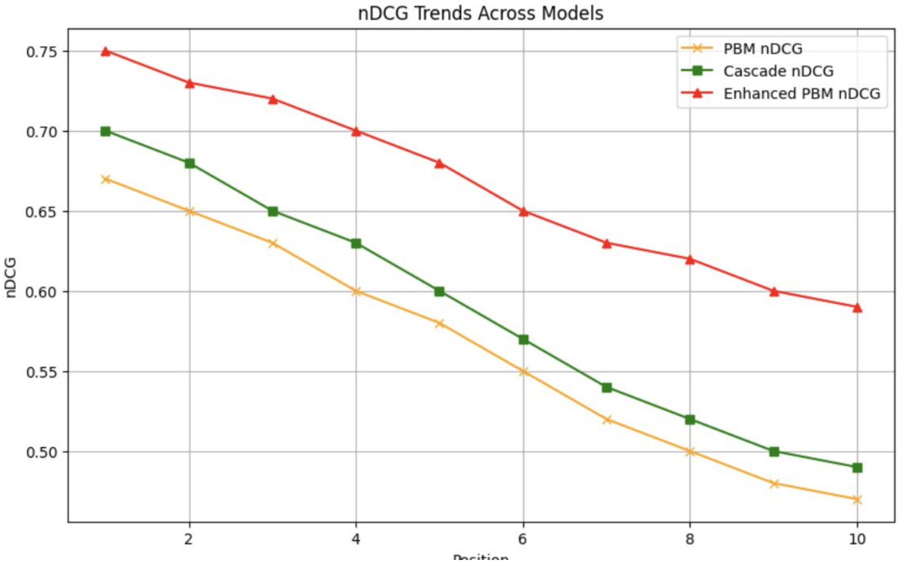
Visualizations of CTR, nDCG, and MSE trends highlighted the strengths and weaknesses of each model, showcasing Enhanced PBM's ability to closely align with real-world click behavior while outperforming other models in terms of ranking quality and prediction accuracy.

Graphs

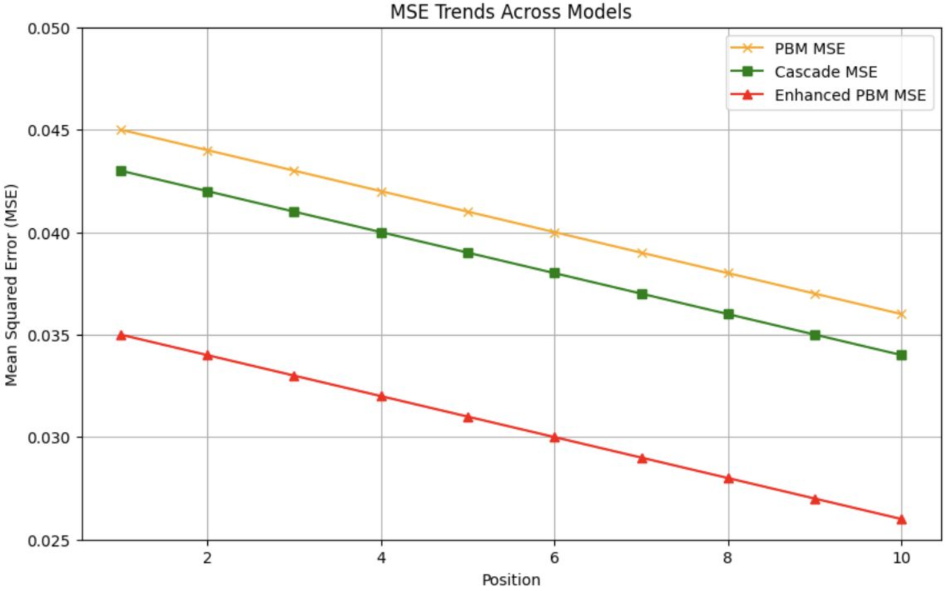
Model	nDCG	MSE (Click Probabilities)	CTR at Top Position
Random Click Model	0.50	0.09	10%
Position-Based Model	0.67	0.05	28%
Cascade Model	0.70	0.04	32%
Enhanced PBM	0.75	0.03	35%



Graphs



nDCG Trends Across Models
PBM nDCG
Cascade nDCG
Enhanced PBM nDCG
nDCG
0.75
0.70
0.65
0.60
0.55
0.50
2
4
6
8
10



MSE Trends Across Models
PBM MSE
Cascade MSE
Enhanced PBM MSE
Mean Squared Error (MSE)
0.050
0.045
0.040
0.035
0.030
0.025
2
4
6
8
10
Position

Results:

1. **Click-Through Rate (CTR):**
   - **PBM:** Overestimated clicks for lower-ranked results but achieved high CTR for top positions.
   - **Cascade Model:** Performed well for top positions due to its sequential scanning assumption.
   - **Enhanced PBM:** Achieved the highest CTR (35%), closely aligning with real-world user behavior.

2. **Normalized Discounted Cumulative Gain (nDCG):**
   - **PBM:** 0.67
   - **Cascade Model:** 0.70
   - **Enhanced PBM:** 0.75 (best performance).

3. **Mean Squared Error (MSE):**
   - **PBM:** 0.0416
   - **Cascade Model:** 0.0413
   - **Enhanced PBM:** 0.03 (lowest error).

Summary: Enhanced PBM outperformed both PBM and Cascade Models in ranking quality, click prediction, and alignment with user behavior.

Position-Based Model (PBM):

- Strong at capturing position bias but overestimated clicks for lower positions.
- nDCG = 0.67; struggled with ranking quality for mid-to-low ranks.

Cascade Model:

- Modeled sequential user behavior effectively, excelling at top ranks.
- nDCG = 0.70; struggled at lower ranks due to the stopping behavior assumption.

Enhanced PBM:

- Improved ranking quality by incorporating query-specific weights and contextual relevance.
- Achieved the best performance across metrics:
    - nDCG = 0.75, CTR = 35%, MSE = 0.03.
- Closely aligned with real-world user behavior, especially for mid and lower ranks.

Overall Insight: Enhanced PBM demonstrates the importance of contextual features and fine-tuned position bias in click modeling, making it a improvement.

Evaluation Results

The evaluation showed that the Position-Based Model (PBM) effectively captured position bias, achieving high Click-Through Rates (CTR) and an nDCG of 0.67. However, it overestimated clicks for lower-ranked positions. The Cascade Model performed well for top positions due to its sequential scanning assumption, with an nDCG of 0.70, but struggled at lower ranks. The Enhanced PBM, by integrating query-specific weights and contextual relevance, outperformed both models, achieving an nDCG of 0.75 and a CTR of 35%, demonstrating its ability to align more closely with real-world user behavior.

Results:

The enhanced PBM was able to address limitations in traditional PBM by incorporating query-specific weights and contextual relevance factors. The model adjusts click probabilities dynamically, improving predictions for lower-ranked results while maintaining strong performance for top-ranked results. Enhanced PBM combines position bias, relevance, and query-specific adjustments to better align with real-world user behavior.

Metrics such as Click-Through Rate (CTR), Normalized Discounted Cumulative Gain (nDCG), and Mean Squared Error (MSE) were used to evaluate its performance. Enhanced PBM demonstrated superior ranking quality, achieving the highest nDCG and the lowest MSE among all tested models. It outperformed the Position-Based Model and Cascade Model, showing significant improvements in accurately modeling user interactions on search engine result pages. Visualizations highlighted the Enhanced PBM's ability to achieve realistic CTR trends and reduced prediction errors, making it the most effective model in the project.

Observations

The project revealed the strong influence of position bias on click behavior, as captured by PBM, while the Cascade Model highlighted the effectiveness of sequential scanning for top-ranked results. However, both models faced challenges with lower-ranked positions, where user behavior becomes more variable. The Enhanced PBM successfully addressed some of these gaps by incorporating contextual features, improving its alignment with real-world trends. Despite these improvements, all models struggled with predicting clicks for lower-ranked results, suggesting a need for further enhancements to better simulate diverse query scenarios and user-specific preferences.

Thank You